

Performance of an Input/Output Buffered Type ATM LAN Switch with Back-Pressure Function

Hiroyuki Ohsaki, Naoki Wakamiya, Masayuki Murata and Hideo Miyahara

Department of Informatics and Mathematical Science
Graduate School of Engineering Science, Osaka University
1-3, Machikaneyama, Toyonaka, Osaka 560, Japan

Tel: +81-6-850-6588

Fax: +81-6-850-6589

E-mail: oosaki@ics.es.osaka-u.ac.jp

Abstract

An ATM switch with both input and output buffers provided with a back-pressure function has been proposed as a cost-effective switch architecture. The back-pressure function prohibits cell transmission from the input buffer to the corresponding output buffer to avoid cell loss at the output buffer due to a temporary congestion. Especially when this switch is applied to ATM LANs for data transfer services, its performance should be evaluated by taking into account bursty traffic. In this paper, we show the maximum throughput, the packet delay distribution, and the approximate packet loss probability of such an ATM switch for bursty traffic through an analytic method. In addition to a balanced traffic condition, an unbalanced traffic and a mixture of bursty and stream traffic are also analyzed. Through several numerical examples, we quantitatively show the effects of the average packet length and the output buffer size on its performance.

Key words: ATM LAN, Input/Output Buffered Type Switch, Back-Pressure Function, Bursty Traffic

1 Introduction

An ATM (Asynchronous Transfer Mode) technology realizes B-ISDN (Broadband Integrated Services Digital Network) by asynchronously treating various kinds of multimedia information such as data, voice and video. The benefit of the ATM technique is enjoyed by a statistical multiplexing of multimedia traffic by dividing it into fixed-size packets (called *cells*). Many efforts of researchers, developments and standardizations have been extensively devoted to public wide area ATM networks. More recently, the ATM technology is also recognized as a promising way for the realization of new high-speed local area networks (LANs) to cope with a rapid advance of high-speed and multimedia-oriented computers (see, e.g., [1]).

Several types of ATM switch architecture such as output buffer switch, input buffer switch, shared buffer switch, batcher banyan switch have been proposed [2, 3]. These switches have trade-offs between performance and implementation complexity. For example, output buffer switch shows better performance than other switches if all switches have a fixed amount of buffer memory. However, since output buffer switch requires memory chips of faster access speed, it cannot be provided with large amount of memory due to cost or technology limitation. As a cost-effective ATM LAN switch, Fan et al. recently proposed a switch architecture that possesses buffers on both sides of input and output ports with a back-pressure function [4]. The key idea of this switch architecture is to provide a large amount of slow-speed (and inexpensive) memory at input ports and a small amount of fast-speed (and expensive) memory at output ports, and to increase its performance by controlling both input and output buffers with back-pressure function. The back-pressure function is provided to avoid a temporary congestion in the switch by prohibiting cell transmission from an input buffer to the congested output buffer when the number of cells in the output buffer exceeds a some threshold value. The performance of this kind of the switch has been analyzed by Iliadis in [5, 6, 7]. However, he assumed that cell interarrival times at each input port follow a geometric distribution. Especially when the above switch is applied to ATM LANs for supporting data transfer service, its performance should be evaluated by taking into account the bursty nature of arriving traffic — packets coming from the upper protocol layers. On the contrary, we will explicitly model such a bursty nature of traffic by assuming that cells (forming a packet) continuously arrive at the input port and are destined for the same output port. More recently, Elwalid et. al have analyzed the performance of multistage switching networks with the back-pressure function for bursty traffic in [8], and Gianatti et. al have analyzed the shared-buffered banyan networks for arbitrary traffic patterns in [9].

However, their target switch architecture is different from ours, and they have treated only cell level performance such as average cell delay and cell loss probability. When an upper layer protocol such as TCP (Transmission Control Protocol) is applied on ATM-based networks, packet (burst) level performance becomes more important. In this paper, we derive the packet delay distribution and the approximate packet loss probability in addition to the maximum throughput.

This paper is organized as follows. In Section 2, we briefly introduce the ATM LAN switch that we will evaluate, and describe its analytic model. In Section 3, the steady state probability of cells in the input/output buffer is first derived. In Section 4, the maximum throughput is then obtained by utilizing the results of Sections 3. The results are extended to an unbalanced traffic condition at each input and output ports, and a mixture with stream traffic as well. In Section 5, we derive the packet delay distribution. In Section 6, the packet loss probability is derived by utilizing a Gaussian approximation. Finally, in Section 7, we conclude our paper with some remarks.

2 Analytic Model

In this section, we describe the ATM LAN switch with back-pressure function followed by an introduction of our analytic model. The number of input ports (and output ports) is represented by N . Our ATM switch is equipped with buffers at both sides of input and output ports (see Fig. 1), and the buffer sizes are denoted by N_I and N_O , respectively. The switching speed of cells from input buffer to output buffer is N times faster than the link speed; that is, in a time slot, at most one cell at the input buffer is transferred to the output buffer, and the output buffer can simultaneously receive N cells from different input buffers. A back-pressure function prohibits transmission of cells from input buffer to output buffer by signaling back from output buffer to input buffer when the number of cells in output buffer exceeds a threshold value [4]. By this control, a cell overflow at output buffer can be avoided (see Fig. 2). However, it introduces HOL (Head Of Line) blocking of cells at input buffer, which results in limitation of the switch performance as will be discussed in the following sections.

We assume that a stream of successively arriving cells forms a packet, and the number of cells in the packet follows a geometric distribution with mean \overline{BL} . Let p denote the probability that at the input port, a newly arriving cell belongs to the same packet. Thus, we have a relation;

$$\overline{BL} = \sum_{i=1}^{\infty} (1-p)ip^{i-1} = \frac{1}{1-p}$$

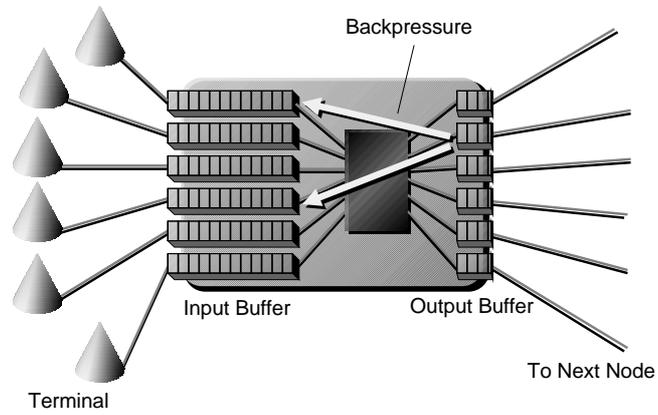


Figure 1: An ATM Switch with Back-Pressure Function.

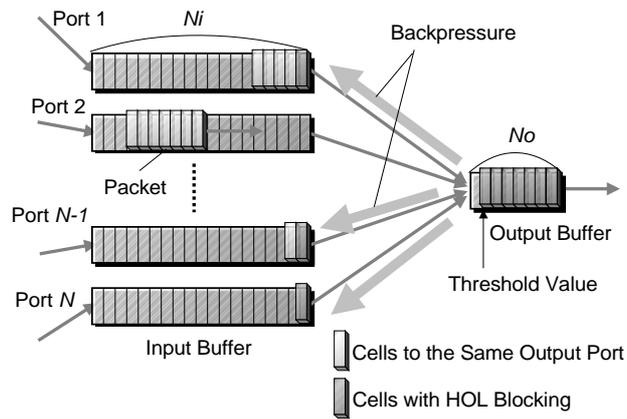


Figure 2: The Analytic Model.

We assume that all cells are stored under first-in-and-first-out (FIFO) discipline at input buffer.

The practical threshold value at output buffer would be $(N_O - N)$ [4]. However, as an ideal case, we assume that the HOL cells are randomly transferred from input buffer to output buffer until the output buffer becomes full. Then, when the output buffer becomes fully occupied, input buffers that have HOL cells destined for this output buffer receives a back-pressure signal to stop cell transmission. Thus, all HOL cells are awaited at the head of input buffers. As soon as the cell in output buffer is transmitted onto the output link, one of HOL cells is selected at random and transmitted to the output buffer. Therefore, it is considered that HOL cells destined for the same output port form a virtual queue, which we will call a HOL queue. While HOL cells are actually stored at the HOL queue, it can be regarded that HOL packets form the HOL queue [5, 6, 7]. Therefore, in what follows, we will use “HOL cell” and “HOL packet” without discrimination.

The switch size N is assumed to be infinity in the following analysis. By introducing this assumption, we can focus on one single output port and its associated HOL queue. The infinite switch size gives the performance limitation as shown in [6, 10]. That is, when compared with the finite case, the maximum throughput with the infinite case gives an upper bound. It is also known that the close values are obtained when N reaches 16 or 32 when the cell interarrivals follow a geometric distribution [6]. In this paper, we will examine this fact even in the case of bursty traffic in Section 4.

In this paper, we will first assume the infinite capacity of the input buffer ($N_I = \infty$) to obtain the maximum throughput (Section 4) and the packet delay distribution (Section 5). Because the memory speed of the output buffer should be N times faster than the link speed, the capacity of the output buffer is limited. On the contrary, the input buffer can operate at the same speed with the input link, i.e., the input buffer can be equipped with large capacity. This assumption is then relaxed to derive the packet loss probability in Section 6. Although the analysis is approximate, it is accurate in the case of the large buffer size as will be validated by comparing with simulation results in Section 6.

3 Derivation of Steady State Probability

We focus on a single output buffer and its associated HOL queue by assuming the infinite number of input and output ports. We consider a discrete-time system where its slot time equals a cell transmission time on the input and output

link. Under the assumptions described in Section 2, the system state is represented by two random variables, Q_k and H_k . Q_k is the number of cells at an output buffer at k th slot and H_k is the number of HOL cells at the input buffers associated with that output buffer. In what follows, the steady state probability of the doublet of two random variables, (Q_k, H_k) , will be derived. For this purpose, we further introduce A_k as a random variable representing the number of HOL packets newly arriving at the HOL queue at the beginning of k th slot. By defining a symbol $(x)^+ = \max(0, x)$, we have the following possibilities.

1. $H_{k-1} + A_k \leq N_O - (Q_{k-1} - 1)^+$; that is, all HOL cells can be transferred to the output port.

At first, we have

$$Q_k = (Q_{k-1} - 1)^+ + H_{k-1} + A_k. \quad (1)$$

Let B_k be the number of the HOL packets that further generate HOL cells at the next $(k + 1)$ th slot. When there exist i HOL packets in the HOL queue, the probability that B_k becomes j is

$$b_{i,j} = \binom{i}{j} p^j (1-p)^{i-j}, \quad (2)$$

and we have

$$H_k = B_k.$$

2. $H_{k-1} + A_k > N_O - (Q_{k-1} - 1)^+$; that is, some HOL cells cannot be transferred to the output port at k th slot.

$(N_O - (Q_{k-1} - 1)^+)$ HOL cells are transferred to the output buffer, and C_k cells of them further generate HOL cells in the next $(k + 1)$ th slot. Therefore, $(H_{k-1} + A_k - (N_O - (Q_{k-1} - 1)^+))$ cells are kept waiting at the HOL queue. Hence, we have

$$\begin{aligned} Q_k &= N_O \\ H_k &= H_{k-1} + A_k - (N_O - (Q_{k-1} - 1)^+) + C_k. \end{aligned}$$

As explained in Section 2, we assume that arrivals of packets at input ports in time slot follow a Poisson distribution since the switch size N is assumed to be infinity. Therefore,

$$a_j \equiv P[A = j] = P[A_k = j] = \frac{\lambda_p^j e^{-\lambda_p}}{j!},$$

where λ_p is the mean arrival rate of packets at each input port. By defining λ_c as the mean arrival rate of cells at input ports, we have

$$\lambda_c = \lambda_p \overline{BL}. \quad (3)$$

We consider $s_{n,m,n',m'}$, which is a transition probability from a state $[Q_{k-1} = n, H_{k-1} = m]$ to $[Q_k = n', H_k = m']$. $s_{n,m,n',m'}$ is obtained as follows.

1. When $n' < N_O$; that is, when the back-pressure function does not work.

From Eq. (1), we have

$$A_k = Q_k - (Q_{k-1} - 1)^+ - H_{k-1}.$$

When m' packets of $(Q_k - (Q_{k-1} - 1)^+)$ HOL packets further generate cells at the next time slot, we have a relation

$$s_{n,m,n',m'} = a_{n'-(n-1)^+-m} b_{n'-(n-1)^+,m'}. \quad (4)$$

2. When $n' = N_O$; that is, when the back-pressure function works.

From Eq. (3), we have

$$A_k = N_O - (Q_{k-1} - 1)^+ - H_{k-1} + (H_k - C_k).$$

Since C_k packets of $(N_O - (Q_{k-1} - 1)^+)$ HOL packets further generate cells at the next time slot, we have

$$s_{n,m,n',m'} = \sum_{i=0}^{m'} a_{n'-(n-1)^+-m+i} b_{n'-(n-1)^+,m'-i}. \quad (5)$$

Let $r_{n,m}$ be the steady state probability defined as

$$r_{n,m} = \lim_{k \rightarrow \infty} P[Q_k = n, H_k = m] = P[Q = n, H = m].$$

In what follows, we will obtain $r_{n,m}$ from Eqs. (4) and (5).

1. When the state is $[Q = 0, H = 0]$, the output port becomes idle. Thus, we have

$$r_{0,0} = 1 - \rho,$$

where ρ is defined as the maximum throughput normalized by the link capacity. By our assumption of the infinite input buffer size, the maximum throughput ρ is equivalent to the cell arrival rate λ_c in steady state if it exists.

2. By considering all states that may change to state $[Q = n - 1, H = 0]$, we have $r_{n,0}$ as follows (see Fig. 3).

$$r_{n,0} = \frac{1}{s_{n,0,n-1,0}} \left\{ r_{n-1,0} - \sum_{i=0}^{n-1} \sum_{j=0}^i s_{i,j,n-1,0} r_{i,j} \right\} \quad (0 < n \leq N_O)$$

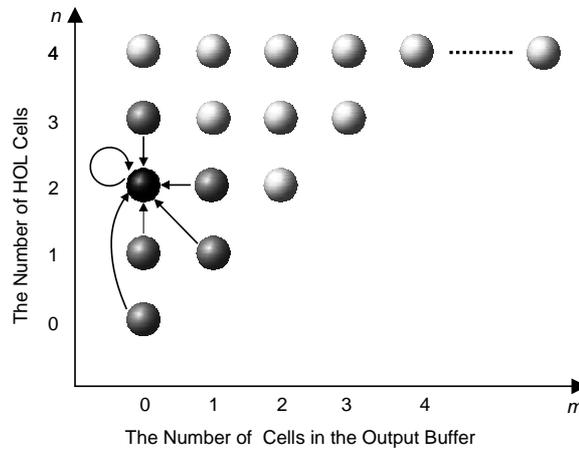


Figure 3: State Transition Diagram in the Case of $m = 0$ and $0 < n \leq N_O$.

3. By considering all states that may change to state $[Q = n, H = m]$, we have $r_{n,m}$ as follows (see Fig. 4).

$$r_{n,m} = \frac{1}{1 - s_{n,m,n,m}} \left\{ \sum_{i=0}^{n-1} \sum_{j=0}^i s_{i,j,n,m} r_{i,j} + \sum_{k=0}^{m-1} s_{n,k,n,m} r_{n,k} \right\} \quad (0 < m, n < N_O)$$

4. By considering all states that may change to the state $[Q = N_O, H = m - 1]$, we have $r_{N_O,m}$ as follows (see Fig. 5).

$$r_{N_O,m} = \frac{1}{s_{N_O,m,N_O,m-1}} \left\{ r_{N_O,m-1} - \sum_{i=0}^{N_O-1} \sum_{j=0}^i s_{i,j,N_O,m-1} r_{i,j} - \sum_{k=0}^{m-1} r_{N_O,k} \right\} \quad (0 < m)$$

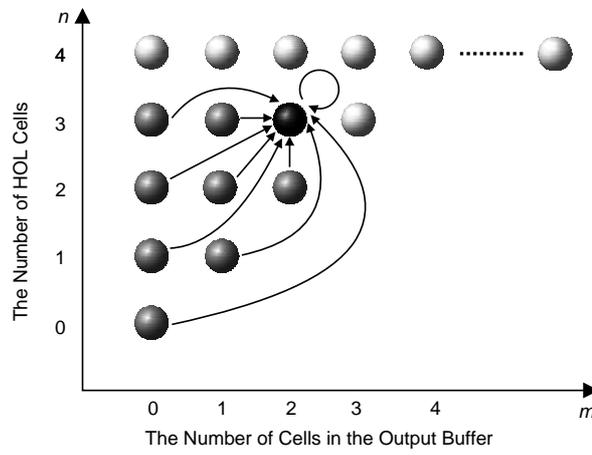


Figure 4: State Transition Diagram in the Case of $0 < m$ and $n < N_0$.

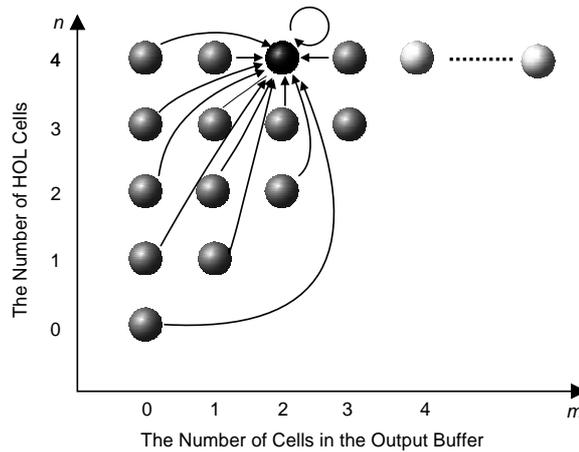


Figure 5: State Transition Diagram in the Case of $0 < m$ and $n = N_0$.

4 Maximum Throughput Analysis

By using the steady state probabilities derived in Section 3, we obtain the maximum throughput under a balanced traffic condition in Subsection 4.1, under an output-unbalanced traffic condition in Subsection 4.2, and under an input-unbalanced traffic condition in Subsection 4.3. The case of a mixture of bursty and stream traffic is also considered in Subsection 4.4.

4.1 Case of Balanced Traffic condition

In this subsection, a balanced traffic condition is assumed; that is, the mean packet arrival rate at every input port is identical and each packet determines its output port with an equal probability $1/N$.

In order to obtain the maximum throughput, we consider the case where all input ports are saturated so that packets are always waiting in HOL queues. In this case, we have a relation:

$$\sum_{i=1}^N A^i = N - \sum_{i=1}^N H^i,$$

where A^i is the random variable which represents the number of arriving packets destined for the output port i in a slot and H^i is the random variable for the number of HOL cells destined for the output port i . By dividing the above equation by N and letting N to be infinity, we have

$$\lambda_p = 1 - \overline{H}, \quad (6)$$

where \overline{H} is the average number of HOL cells. \overline{H} is expressed with $r_{n,m}$ derived in Section 3 as

$$\overline{H} = \sum_{n=0}^{N_O} \sum_{m=1}^{\infty} m r_{n,m}.$$

From Eqs. (3) and (6), we have

$$\lambda_c = (1 - \overline{H})\overline{BL}. \quad (7)$$

The maximum throughput ρ can be obtained by substituting λ_c in the above equation with ρ and solving it for ρ . Since \overline{H} depends on ρ , ρ is solved iteratively by virtue of a standard iteration technique such as a bisection method [11].

In Figs. 6 and 7, the maximum throughput ρ is plotted for the average packet length \overline{BL} and the output buffer size N_O , respectively. These figures show that the packet length drastically degrades the maximum throughput. Furthermore, we may observe that the size of output buffers must be larger than the average packet length to gain a sufficient throughput. We note that the maximum throughput for $N_O = 1$ is exactly same as the well known value of the input queuing, 0.585 [10]. Figure 8 compares the analytic results (the switch size $N = \infty$) with simulation

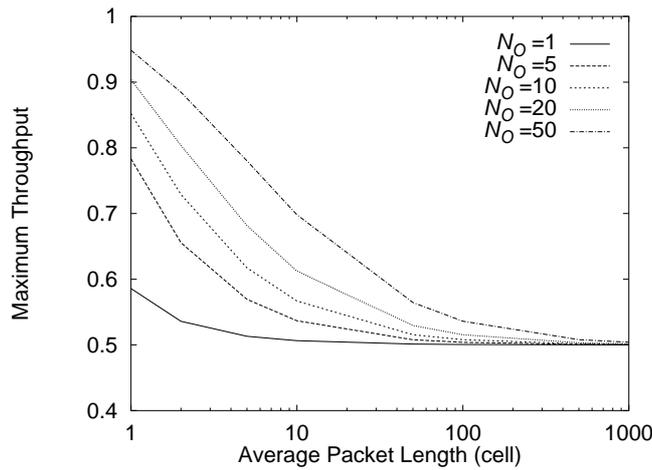


Figure 6: Maximum Throughput vs. Average Packet Length.

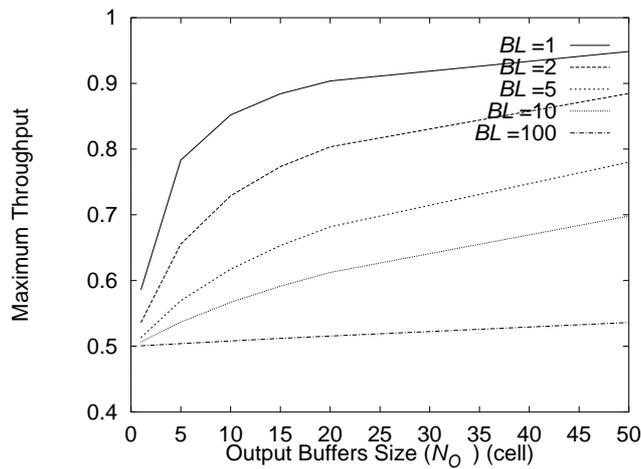


Figure 7: Maximum Throughput vs. Output Buffer Size.

results ($N = 8, 16$ and 32) for $N_O = 1$ and $N_O = 50$ dependent on the average packet length \overline{BL} . The case of

$\overline{BL} = 1$ in the figure corresponds to results obtained in [5, 6, 7], and it can be found that the analytic results become close to simulation results as the switch size gets large even in the case of $BL > 1$. We note that 95% confidence intervals of all simulation results for maximum throughput are within 2% of mean values, and are not shown in the figure.

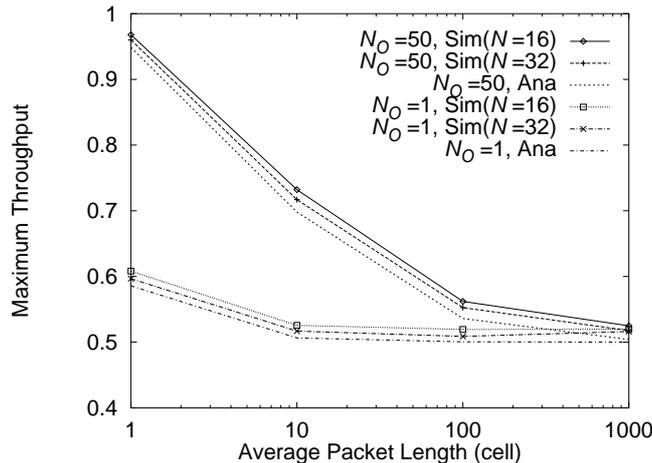


Figure 8: Comparison with Simulation Results.

4.2 Case of Unbalanced Traffic at Output Ports

In this section, output unbalanced traffic is treated following the approach presented in [5]. Output buffers are divided into two groups called O_1 and O_2 . Let q_O be a ratio of the number of output ports belonging to the group O_1 as

$$q_O \equiv \frac{|O_1|}{N}. \quad (8)$$

The packet arrival rate at each input port is identical. However, each packet arriving at the input port selects one of output ports in group O_1 with probability P_{G1} or one of output ports in group O_2 with probability P_{G2} . By assuming $P_{G1} \geq P_{G2}$ without loss of generality, the relative probability r_O is denoted as

$$r_O \equiv \frac{P_{G1}}{P_{G1} + P_{G2}} \geq 0.5. \quad (9)$$

It is noted that the balanced traffic case is a special case by setting $q_O = 0$, $q_O = 1$ or $r_O = 0.5$. Let P_1 and P_2 be the probabilities that an arriving packet is destined to output ports belonging to the O_1 and O_2 , respectively. From

Eqs. (8) and (9), we have

$$\begin{aligned}
 P_1 &= \frac{|O_1|P_{G1}}{|O_1|P_{G1} + (N - |O_1|)P_{G2}} \\
 &= \frac{\varrho r_O}{1 - \varrho - r_O + 2\varrho r_O} \\
 P_2 &= \frac{(N - |O_1|)P_{G2}}{|O_1|P_{G1} + (N - |O_1|)P_{G2}} \\
 &= \frac{1 - \varrho - r_O + \varrho r_O}{1 - \varrho - r_O + 2\varrho r_O}.
 \end{aligned}$$

We define λ_p as the packet arrival rate at each input port, and λ_{p1} and λ_{p2} as the packet arrival rates at output ports belonging to the group O_1 and O_2 , respectively. We then obtain

$$\begin{aligned}
 \lambda_{p1} &= \frac{r_O \lambda_p}{1 - \varrho - r_O + 2\varrho r_O} \\
 \lambda_{p2} &= \frac{(1 - r_O) \lambda_p}{1 - \varrho - r_O + 2\varrho r_O}.
 \end{aligned}$$

For deriving the maximum throughput, we consider a relation

$$\sum_{i=1}^N A^i = N - \left(\sum_{i=1}^{|O_1|} H_1^i + \sum_{i=1}^{|O_2|} H_2^i \right),$$

where random variables H_1^i (H_2^i) is the number of HOL cells destined for the output port belonging to the group O_1 (O_2). By dividing the above equation by N and letting N to be infinity, we have

$$\lambda_p = 1 - \{\varrho \bar{H}_1 + (1 - \varrho) \bar{H}_2\},$$

where \bar{H}_1 and \bar{H}_2 are the average number of HOL cells destined for the group O_1 and O_2 , respectively. From Eq. (3), we have

$$\lambda_c = \left[1 - \{\varrho \bar{H}_1 + (1 - \varrho) \bar{H}_2\} \right] \bar{BL}.$$

The maximum throughput ρ can be obtained by substituting λ_c with ρ in the above equation and solving for ρ in the same manner presented in Subsection 4.1.

In Figs. 9 and 10, the relations between ϱ and the maximum throughput are plotted for $\bar{BL} = 1$ and $\bar{BL} = 10$, respectively. These figures show that an unbalanced traffic and a larger packet size cause degradation of the maximum throughput.

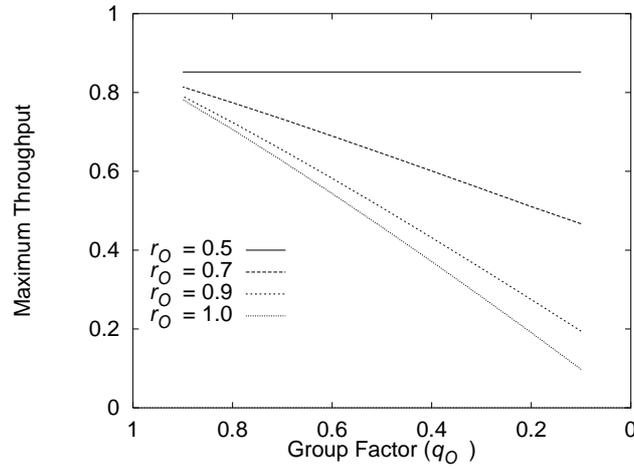


Figure 9: Unbalanced Traffic at Output Ports ($N_O = 10$ and $\overline{BL} = 1$).

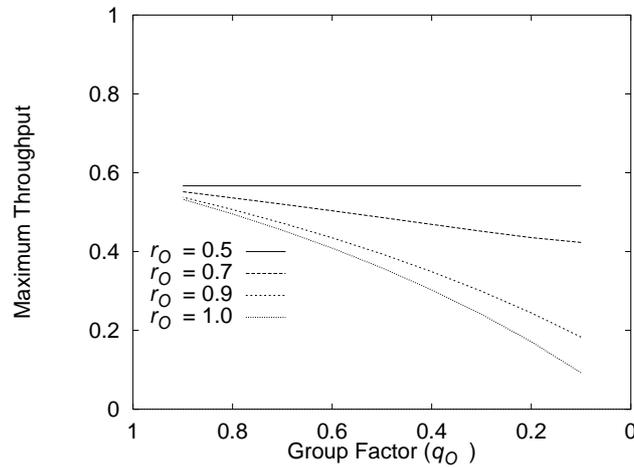


Figure 10: Unbalanced Traffic at Output Ports ($N_O = 10$ and $\overline{BL} = 10$).

4.3 Case of Unbalanced Traffic at Input Ports

In this subsection, we evaluate the performance of the switch under an unbalanced traffic condition at the input ports. Similar to the previous subsection, input ports are divided into two groups I_1 and I_2 . Let q_I be a ratio of the number of input ports belonging to the group I_1 defined as

$$q_I \equiv \frac{|I_1|}{N}. \quad (10)$$

We further introduce λ_{p1} and λ_{p2} as mean packet arrival rates at the groups I_1 and I_2 , respectively. Assuming that $\lambda_{p1} \geq \lambda_{p2}$ without loss of generality, we introduce r_I as

$$r_I \equiv \frac{\lambda_{p1}}{\lambda_{p1} + \lambda_{p2}} \geq 0.5. \quad (11)$$

It is noted that the balanced traffic case is the special case by setting $q_I = 0$, $q_I = 1$ or $r_I = 0.5$. We assume that each packet arriving at the input port chooses the output port with a same probability $1/N$. By letting λ_p denote the packet arrival rate at each output port, λ_{p1} and λ_{p2} are given as

$$\begin{aligned} \lambda_{p1} &= \frac{\lambda_p r_I}{1 - q_I - r_I + 2q_I r_I} \\ \lambda_{p2} &= \frac{\lambda_p (1 - r_I)}{1 - q_I - r_I + 2q_I r_I}. \end{aligned}$$

To obtain the maximum throughput, we consider the case where input ports are saturated. Recalling that we assume $\lambda_{p1} \geq \lambda_{p2}$, the input buffers belonging to the group I_1 is saturated first. Thus, we have

$$\sum_{i=1}^{|I_1|} A_1^i = |I_1| - \sum_{i=1}^N \frac{\lambda_{p1}}{\lambda_p} \frac{|I_1|}{N} H^i,$$

where the random variable A_1^i is the number of packets arriving at the input port i belonging to the group I_1 . By dividing the above equation by N and letting N to be infinity, we have

$$\lambda_{p1} = 1 - \frac{r_I \bar{H}}{1 - q_I - r_I + 2q_I r_I}.$$

From Eq. (3), we obtain

$$\lambda_{c1} = \left(1 - \frac{r_I \bar{H}}{1 - q_I - r_I + 2q_I r_I}\right) \bar{B}\bar{L},$$

where λ_{c1} is the mean packet arrival rate at each input port belonging to the group I_1 . The maximum throughput ρ can be obtained by substituting λ_{c1} in the above equation with ρ and solving for ρ as in the same manner presented in Subsection 4.1.

Figures 11 and 12 show the maximum throughput dependent on q for $\overline{BL} = 1$ and $\overline{BL} = 10$, respectively. These figures show that an unbalanced traffic condition and a larger packet size degrade the maximum throughput. The result for $\overline{BL} = 1$ is almost same as that for the output unbalanced traffic (Fig. 9). On the other hand, the result for $\overline{BL} = 10$ shows higher performance than that of output unbalanced traffic (Fig. 10). This is because unbalanced traffic at input ports causes less HOL blocking than at output ports.

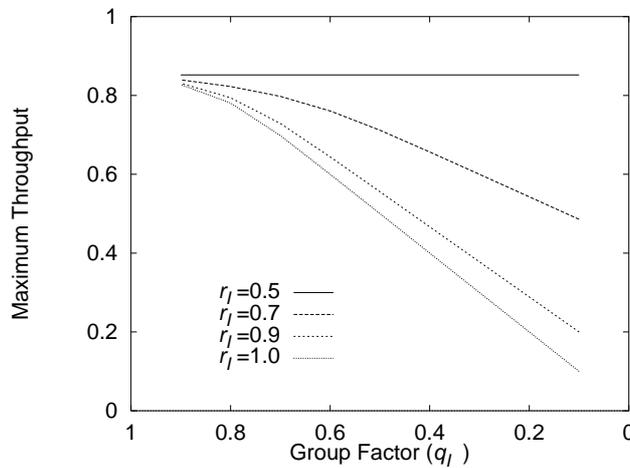


Figure 11: Unbalanced Traffic at Input Ports ($N_O = 10$ and $\overline{BL} = 1$).

4.4 Case of Mixture with Stream Traffic

Finally, we derive the maximum throughput in the case where the bursty traffic and the stream traffic coexist. We assume that the stream traffic occupies some portion of the link with constant peak rate. For example, this class of traffic can support an uncompressed video transfer service.

Let R denote the peak rate of stream traffic normalized by the link capacity. The switch can simultaneously accept $m(\leq \lfloor 1/R \rfloor)$ calls of stream traffic. We assume that call arrivals of the stream traffic follow a Poisson distribution with mean λ_{CBR} , and its service time (call holding time) has an exponential distribution with mean $1/\mu_{CBR}$. While both bursty and stream traffic share a link, cells of the stream traffic are given higher priority. Namely, cells

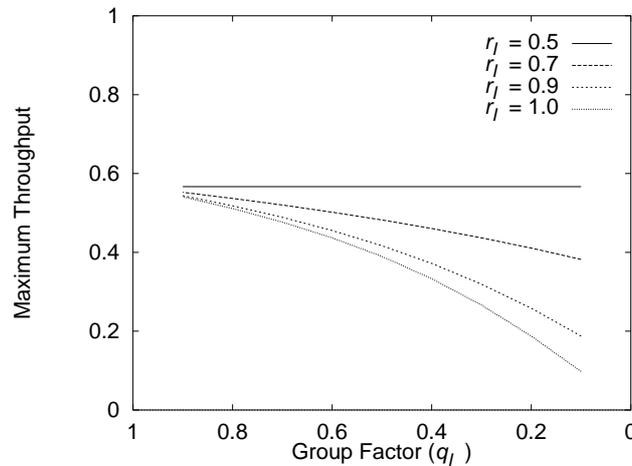


Figure 12: Unbalanced Traffic at Input Ports ($N_O = 10$ and $\overline{BL} = 10$).

of stream traffic arriving at the input port are transferred to its destination output port prior to cells of bursty traffic [4]. By this control mechanism, it can be considered that bursty traffic can utilize $1 - nR$ of the link capacity when n calls of stream traffic are accepted. We note that if compressed video transfer service is accommodated as stream traffic, more capacity can be utilized by bursty traffic. Thus, the maximum throughput derived in the below should be regarded as the “minimum” guaranteed throughput for the bursty traffic.

Since the stream traffic is given high priority, it can be modeled by an M/M/m/m queuing system. By letting π_n be the probability that n calls of stream traffic are accepted in steady state, π_n is given as follows (e.g., [12]).

$$\pi_n = \left[\sum_{n=0}^m \left(\frac{\lambda_{CBR}}{\mu_{CBR}} \right)^n \frac{1}{n!} \right]^{-1} \left(\frac{\lambda_{CBR}}{\mu_{CBR}} \right)^n \frac{1}{n!}$$

Since the service time of steam traffic can be assumed to be much longer than cell or the packet transmission time of bursty traffic, the available link capacity for bursty traffic is regarded to be constant when the number of accepted calls of stream traffic is fixed. By letting ρ_n be the maximum throughput for bursty traffic when n calls of the stream traffic are accepted, we have [4]

$$\rho_n = (1 - nR)\rho,$$

where ρ is defined as the maximum throughput of bursty traffic when all link capacity is allocated to bursty traffic, and has been already derived in Subsection 4.1. Consequently, the “averaged” maximum throughput ρ is obtained

as follows.

$$\rho' = \sum_{n=0}^m \pi_n \rho_n$$

Figure 13 shows the maximum throughput of bursty traffic and throughput of stream traffic dependent on an offered traffic load for stream traffic for $N_O = 50$, $\mu_{CBB} = 0.1$, $R = 0.2$ and $m = 5$. From this figure, we can observe the natural idea that the larger the average packet length is, the smaller the maximum allowable throughput of bursty traffic is. Therefore, the available bandwidth allocated to the stream traffic should be limited in some way to avoid a degradation of bursty traffic efficiency. One possible approach is to decrease m , which is the maximum number of calls of stream traffic that the switch can accept. In an actual situation, it can be implemented in CAC (Call Admission Control) so that an acceptable number of calls of stream traffic is limited. Figure 14 shows the maximum throughput of both bursty traffic and stream traffic dependent on the offered traffic load for stream traffic for $\overline{BL} = 1$ and several values of m . It shows that the performance degradation of bursty traffic can be avoided to some extent by limiting m .

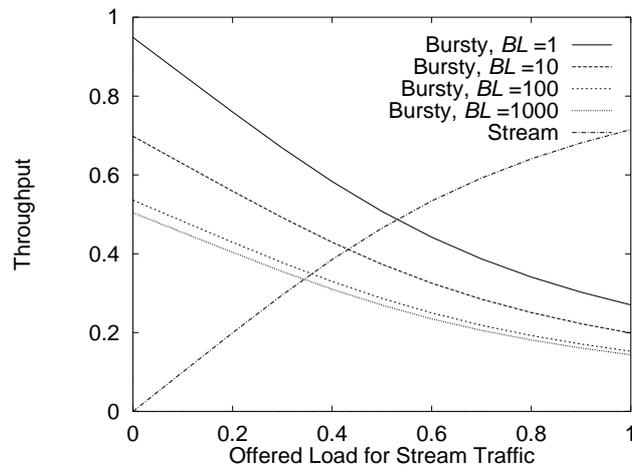


Figure 13: Throughput vs. Offered Load for Stream Traffic.

5 Derivation of Packet Delay Distribution

In this section, we derive the packet delay experienced at both input and output buffer. The packet delay is defined as the time duration from when the first cell of the packet arrives at the input port of the switch to when the last cell

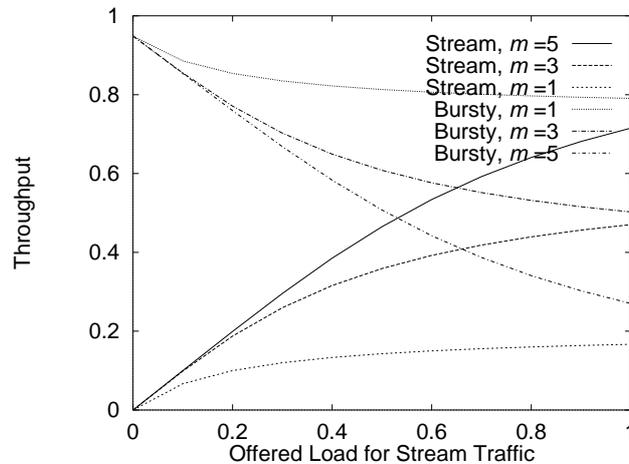


Figure 14: Effect of Available Link Capacity Limitation on Stream Traffic.

is transmitted onto the output link. We divide the packet delay into the following three elements.

1. W_I : The packet waiting time at the input buffer from the arrival time of the first cell of the packet at the input buffer to its arrival time at the HOL queue.
2. W_S : The switching delay from the HOL queue to its destination output port; that is, the time duration from the arrival time of the first cell at the HOL queue to the departure time of the last cell from the HOL queue.
3. W_O : The packet waiting time at the output buffer from the arrival time of the first cell of the packet at the output buffer to its departure time from the output buffer.

It is assumed that the cell transmission from the HOL queue is performed by a random discipline for cells arriving in the same slot, and by a FIFO discipline for cells arriving in different slots. In Subsections 5.1, 5.2 and 5.3, we will derive the above three elements.

5.1 Switching Delay

For obtaining the switching delay W_S , we examine the cell transmission behavior of the tagged packet arriving at the HOL queue. Let u_m be the probability that the number of packets waiting in the HOL queue including the just

arriving tagged packet equals m , which is obtained as

$$u_m = \sum_{n=0}^{N_O} \sum_{j=1}^m r_{n,m-j} a'_j,$$

where a'_j is the probability that the tagged packet arrives with j packets in the same slot; that is,

$$a'_j = \frac{j a_j}{\sum_{k=1}^{\infty} k a_k} = \frac{j a_j}{\lambda_p}.$$

In what follows, we will refer a cycle to the time to transfer all cells of the tagged packet from the HOL queue to the output buffer.

Suppose now that there are m packets including the tagged one in the HOL queue at the beginning of the cycle, that j packets of them have more cells to transfer, and that $(m' - 1 - j)$ packets newly arrive at the HOL queue during the cycle. In this case, the transition probability $t_{m,m'}$ is given as

$$t_{m,m'} = \sum_{j=0}^{m'-1} b_{m-1,j} a_{m'-1-j}^m,$$

where a_k^m is defined as the probability that k packets arrives at the HOL queue during m slots; that is,

$$a_k^m = \frac{(\lambda_p m)^k e^{-\lambda_p m}}{k!}.$$

Let $T_{m,m'}(k)$ be the cycle time distribution when m HOL cells exist at the beginning of the cycle, and when there are m' HOL cells at the beginning of the next cycle. Using the above probability $t_{m,m'}$, $T_{m,m'}(k)$ is expressed as follows.

$$T_{m,m'}(k) = \begin{cases} t_{m,m'}, & \text{if } k = m \\ 0, & \text{otherwise} \end{cases}.$$

By letting $T_{m,m'}^l(k)$ be the distribution over l cycles, we have

$$T_{m,m'}^l(k) = \sum_{j=1}^{\infty} [T_{m,j}^{l-1} \otimes T_{j,m'}](k),$$

where the symbol \otimes is the convolution operator of two probability distributions; that is, for two probability distributions $y_1(k)$ and $y_2(k)$, it is defined as

$$[y_1 \otimes y_2](k) \equiv \sum_{j=0}^k y_1(j) y_2(k-j).$$

Next, let $U_m(k)$ represents the delay distribution of the last cell of the tagged packet. Because of our assumption that the cell transmissions are done by a random discipline among cells arriving at the HOL queue in the same slot, we have

$$U_m(k) = \begin{cases} 1/m, & \text{if } 0 \leq k \leq m-1 \\ 0, & \text{otherwise} \end{cases}.$$

We further introduce $W_m(k)$ that is denoted as the transmission time distribution of the tagged packet conditioned on m , which is the number of HOL packets when the tagged packet arrives at the HOL queue. Recalling that the packet length (the number of cells in the packet) follows a geometric distribution with parameter p , $W_m(k)$ is given by

$$W_m(k) = (1-p)U_m(k) + \sum_{l=1}^{\infty} p^l (1-p) \sum_{j=1}^{\infty} [T_{m,j}^l \otimes U_j](k).$$

Hence, the mean switching delay W_S is obtained as

$$W_S = \sum_{m=1}^{\infty} \sum_{k=1}^{\infty} kW_m(k)u_m.$$

5.2 Packet Waiting Time at Input Buffer

In order to obtain W_I , we first consider the random variable W_H , the time from when the first cell of the packet arrives at the HOL queue to when all cells belonging to the same packet are transferred to the output buffer. The derivation of distribution for W_H is similar to that of W_S , but in addition to the state of the HOL queue, the state of the output buffer should be taken into account. Let $u_{n,m}$ be the probability that there are m packets in the HOL queue and n cells in the output buffer at the arriving instant of the tagged packet. It is determined as

$$u_{n,m} = \begin{cases} \sum_{j=1}^m (r_{0,m-j} + r_{1,m-j})a'_j, & \text{if } n = 0 \\ \sum_{j=1}^m r_{n+1,m-j}a'_j, & \text{otherwise} \end{cases}$$

We define $C_{n,m,n',m'}(k)$ as the probability distribution of a cycle time that the state was (n, m) at the beginning of a cycle, and that the state becomes (n', m') at the beginning of the next cycle. It is noted that the current definition of the cycle is different from that in the previous subsection in the sense that it is observed at the HOL queue. More precisely, when the output buffer has space to accept, say three cells, three cells can be transmitted simultaneously in one slot from the HOL queue if those exist, and in the current definition of the cycle, it is counted as one slot. On

the other hand, in the previous subsection, it is counted as three slots to derive the switching delay. $C_{n,m,n',m'}(k)$ is obtained dependent on m and n as follows.

- $m \leq N_O - n$

Since all HOL cells can be transferred to the output buffer, the cycle time is just one slot. The state of the output buffer then becomes $n' = n + m$. On the other hand, the number of HOL packets becomes $m' = j + k + 1$ when j of HOL packets (except the tagged one) have more cells to transfer and when k packets newly arrive in the current cycle. Consequently, we have

$$C_{n,m,n',m'}(k) = \begin{cases} \sum_{j=0}^{m'} b_{m-1,j} a_{m'-1-j}, & \text{if } k = 1 \text{ and } n' = n + m \\ 0, & \text{otherwise} \end{cases}.$$

- $m > N_O - n$

$(N_O - n)$ cells are transferred to the output buffer in one slot, and the other $(m - (N_O - n))$ cells are transferred continuously in the following slots. Therefore, the cycle time is $(1 + m - (N_O - n))$, and the state of the output buffer becomes $n' = N_O$. When j packets of $(m - 1)$ HOL packets have more cells to transfer and when k packets arrive at the current cycle, the number of HOL packets becomes $m' = k + 1$. Therefore, we have

$$C_{n,m,n',m'}(k) = \begin{cases} \sum_{j=0}^{m'} b_{m-1,j} a_{m'-1-j}^{m-(N_O-n)+1}, & \text{if } k = m - (N_O - n) \text{ and } n' = N_O \\ 0, & \text{otherwise} \end{cases}.$$

The cycle time distribution over l cycles is then obtained as

$$C_{n,m,n',m'}^l(k) = \sum_{n''=0}^{N_O} \sum_{m''=1}^{\infty} [C_{n,m,n'',m''}^{l-1} \otimes C_{n'',m'',n',m'}](k).$$

Let $U_{n,m}(k)$ be the delay distribution of the last cell of the packet in the cycle. Because of our assumption that the cell transmission is done by a FIFO discipline among cells arriving in distinct slots, $U_{n,m}(k)$ is given as follows.

- $m \leq N_O - n$

$$U_{n,m}(k) = \begin{cases} 1, & \text{if } k = 0 \\ 0, & \text{otherwise} \end{cases}$$

- $m > N_O - n$

$$U_{n,m}(k) = \begin{cases} (N_O - n)/m, & \text{if } k = 0 \\ 1/m, & \text{if } k \leq m - (N_O - n) \\ 0, & \text{otherwise} \end{cases}$$

Probability distribution of W_H is obtained as

$$W_H(k) = \sum_{n=0}^{N_O} \sum_{m=1}^{\infty} u_{n,m} \left[(1-p)U_{n,m}(k) + \sum_{l=1}^{\infty} p^l (1-p) \sum_{n'=0}^{N_O} \sum_{m'=1}^{\infty} [C_{n,m,n',m'}^l \otimes U_{n',m'}](k) \right].$$

The corresponding n th moment $W_H^{(n)}$ is then given by

$$W_H^{(n)} = \sum_{k=1}^{\infty} k^n W_H(k).$$

Finally, by considering a Geom/G/1 queuing system where the first and second moments of the service time are given by $W_H^{(1)}$ and $W_H^{(2)}$, respectively, we have (see, e.g., [13])

$$W_I = \frac{\lambda_p W_H^{(2)}}{2(1 - \lambda_p W_H^{(1)})}.$$

5.3 Packet Waiting Time at Output Buffer

Since W_O means the delay of the first cell of the packet in the output buffer, we simply have

$$W_O = 1 + \sum_{n=1}^{N_O} \sum_{m=0}^{\infty} n r_{n,m},$$

which includes the transmission time of the last cell.

5.4 Numerical Examples

Figures 15 and 16 show relations between the offered load and the average packet delay for $\overline{BL} = 1$ and $\overline{BL} = 3$, respectively, for various values of output buffer size N_O . In Fig. 16, simulation results for the switch size $N = 16$ are also provided due to computational complexity of our analytic approach. In simulation, we have set the switch size N to 16 in obtaining the results for larger N_O 's. These figures show that the high offered load suddenly increases the average packet delays, which becomes saturated at the point where the offered load reaches the maximum throughput. Inversely, if we use an appropriate size of the output buffer, it would be possible to sustain increase of the

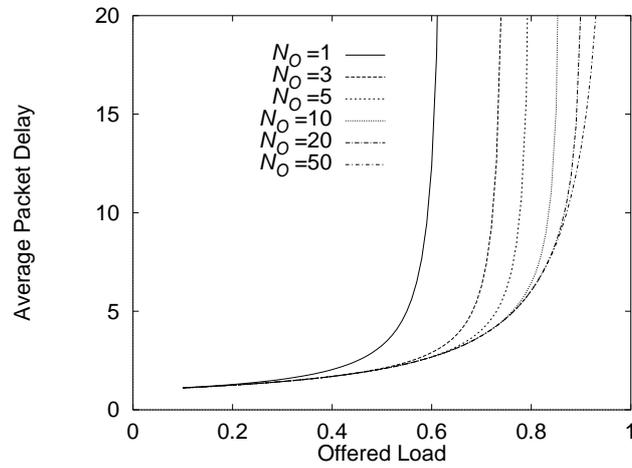


Figure 15: Average Packet Delay vs. Offered Load for $\overline{BL} = 1$.

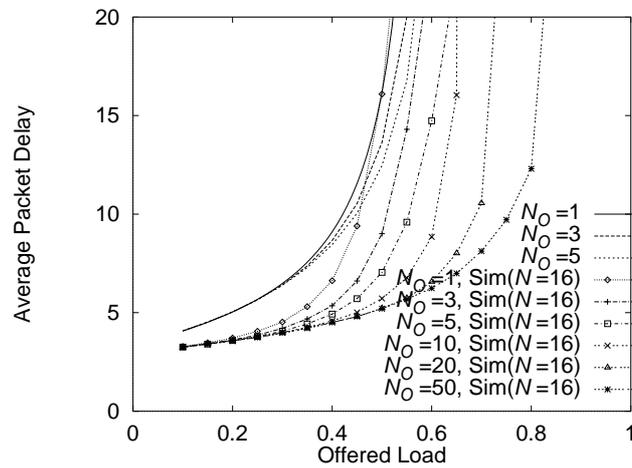


Figure 16: Average Packet Delay vs. Offered Load for $\overline{BL} = 3$.

average packet delay as having shown in Subsection 4 (see Fig. 6), but it is limited as the mean packet length becomes large. To validate our analytic method, we provide simulation results as well as analytic ones in Fig. 17 for $N_O = 5$ and $\overline{BL} = 1$ and $\overline{BL} = 3$. It can be found that our analysis gives slightly larger value than simulation. It is just because the switch size N is assumed to be infinite in our analysis.

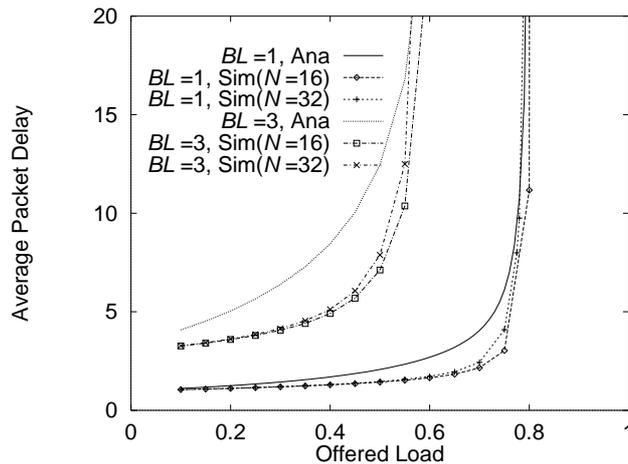


Figure 17: Comparison with Simulation Results.

6 Approximate Analysis of Packet Loss Probability

In this section, the packet loss probability is derived by utilizing a Gaussian approximation. In addition to the FIFO switch considered above, the RIRO (Random-In-Random-Out) switch [4] is also considered for comparison. In the RIRO switch, in order to avoid the HOL blocking, all cells at each input buffer are stored in logically separated buffers, each of which is associated with the destination output port. The packet loss probabilities for these two switches are approximately derived in the followings.

6.1 Case of FIFO Switch

At first, we consider a discrete time Geom/G/1 queuing system where packet interarrival times follow a geometric distribution with parameter λ_p . We define $\Lambda(z)$ as the probability generation function (PGF) for the distribution of

the number of packets arriving in a slot, which is given by

$$\Lambda(z) = 1 - \lambda_p + \lambda_p z.$$

Furthermore, we let $B(z)$ be the PGF of probability distribution of the service time of the customers. Its i th derivative is defined by $b^{(i)}$; that is,

$$b^{(i)} \equiv \left. \frac{d^i B(z)}{dz^i} \right|_{z=1}.$$

The PGF of the unfinished work for this system is given by (see, e.g., [13])

$$U(z) = \frac{(1 - \rho)(1 - z)\Lambda[B(z)]}{\Lambda[B(z)] - z},$$

where ρ is the utilization obtained as

$$\rho = \lambda_p b^{(1)}.$$

The average and the variance of $U(z)$ is derived as

$$\begin{aligned} E[U] &= \left. \frac{dU(z)}{dz} \right|_{z=1} \\ &= \frac{\lambda_p b^{(2)} + \lambda_p^{(2)} (b^{(1)})^2 + \rho(1 - 2\rho)}{2(1 - \rho)} \\ V[U] &= E[U^2] - E[U]^2, \end{aligned}$$

where $E[U^2]$ is given by

$$E[U^2] = \left. \frac{d^2 U(z)}{dz^2} \right|_{z=1} + E[U]^2$$

In the FIFO switch, we can view the number of cells in the input buffer as the unfinished work. Therefore, the packet loss probability P_L is approximately given as

$$P_L(FIFO) \cong Pr[U > N_I] = \int_{N_I}^{\infty} \frac{1}{\sqrt{2\pi V[U]}} e^{-\frac{(y - E[U])^2}{2V[U]}} dy, \quad (12)$$

where N_I represents the buffer size. The probability distribution of W_H obtained in Section 5 can be applied to Eq. (12) for the moments of the service time distribution. Namely, $b^{(i)}$'s ($1 \leq i \leq 3$) are given by

$$b^{(1)} = W_H^{(1)} \quad (13)$$

$$b^{(2)} = W_H^{(2)} - W_H^{(1)} \quad (14)$$

$$b^{(3)} = W_H^{(3)} - 3W_H^{(2)} + 2W_H^{(1)}. \quad (15)$$

6.2 Case of RIRO Switch

We assume that each input buffer is composed of N Geom/G/1 queues, each of which is associated with the output port. We further assume that each queue is served independently. This assumption is realistic if the switch performs an appropriate cell transmission scheduling [4]. Furthermore, by assuming balanced traffic load condition, the mean packet arrival rate at the j th queue at the input buffer (dedicated to the output port j) is given as

$$\lambda_j = \frac{\lambda_p}{N}.$$

By letting $\Lambda_j(z)$ be the z -transform for the number of packets arriving in a slot, we have

$$\Lambda_j(z) = 1 - \lambda_j + \lambda_j z.$$

We define V_j as a random variable for the number of cells waiting at the j th queue in the input buffer. To prevent a single queue from occupying the whole input buffer, the threshold value T_h is introduced for all queues, and the packet loss probability due to this threshold value T_h is given by

$$P(T_h) \cong Pr[V_j > T_h]$$

The packet service time distributions for each queue are obtained from Eq. (15) by letting λ be λ_j .

Next, let U_N be the random variable to represent the unfinished work defined as

$$U_N = \sum_{j=1}^N V_j.$$

By introducing $U_N(z) = V_j(z)^N$ for the PGF of U_N , the average and the variance of U_N are obtained as follows.

$$\begin{aligned} E[U_N] &= \left. \frac{dV_j^N(z)}{dz} \right|_{z=1} \\ V[U_N] &= E[U_N^2] - E[U_N]^2 \end{aligned}$$

Let P_L denote the probability that the number of cells at the input buffer exceeds the physical buffer size N_I , we have

$$P_L(RIRO) \cong Pr[\lim_{N \rightarrow \infty} \sum_{j=1}^N V_j > N_I] = Pr[\lim_{N \rightarrow \infty} U_N > N_I].$$

Consequently, the packet loss probability for the RIRO switch, $P(RIRO)$, is obtained as

$$P(RIRO) \cong \max(P(T_h), P_L(RIRO)).$$

6.3 Numerical Examples

In Figs. 18 and 19, packet loss probabilities dependent on the offered load are plotted for $\overline{BL} = 1$ and $\overline{BL} = 3$, respectively. For comparison purposes, we also provide the result of the output buffer switch [10]. Here, we set $N_I + N_O = 30$ in the cases of FIFO and RIRO switches and $N_O = 30$ in the case of the output buffer switch. In both cases of FIFO and RIRO switches, the higher offered load results in sudden degradation of the packet loss probability. The FIFO switch gives the larger packet loss probability than both of the RIRO switch and the output buffer switch for the same buffer size. However, the performance of the FIFO switch can be further improved by a large capacity of the input buffer with low speed memory while the output buffer switch requires high speed buffers at the output ports. It should be noted from Fig. 18 that FIFO switch with $N_O = 5$ shows better performance than RIRO switch with $N_O = 1$ when the offered load is rather low. This is explained as follows. Performance degradation of FIFO switch is mainly caused by HOL blocking. However, when the offered load is much lower than the maximum throughput, HOL blocking rarely occurs. Thus, FIFO switch with larger output buffer ($N_O = 5$) gains lower packet loss probability than RIRO switch with smaller output buffer ($N_O = 1$). Of course, if the same amounts of buffers are equipped at input/output buffers, RIRO switch gives higher performance than FIFO switch at the expense of more complicated hardware implementation. Furthermore, RIRO switch gives lower packet loss probabilities than the output buffer switch even though it requires a less amount of high-speed output buffer memory.

Finally, we assess the accuracy of our analytic results by comparing with simulation results. Figures 20 and 21 illustrate the comparison results for the packet length $\overline{BL} = 1$ and $\overline{BL} = 3$, respectively, for the FIFO switch. Since our approach is based on the Gaussian approximation method, only the small packet loss probabilities are meaningful as indicated in the figures.

7 Conclusion

In this paper, an ATM switch with input and output buffer equipped with back-pressure function was treated. We have analyzed its performance under bursty traffic condition for applying it to ATM LANs. We have derived the maximum throughput and the packet delay distribution as well as the approximate packet loss probability under the assumption that the switch size is infinite. Consequently, we have shown that larger packet lengths drastically degrade the performance of the switch. However, it is possible to sustain such a degradation to some extent by

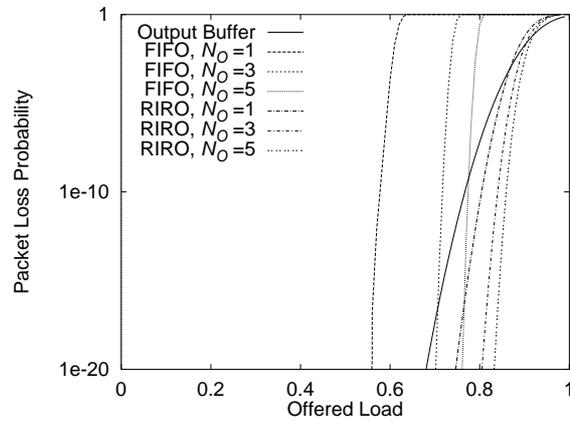


Figure 18: Packet Loss Probability vs. Offered Load for $\overline{BL} = 1$.

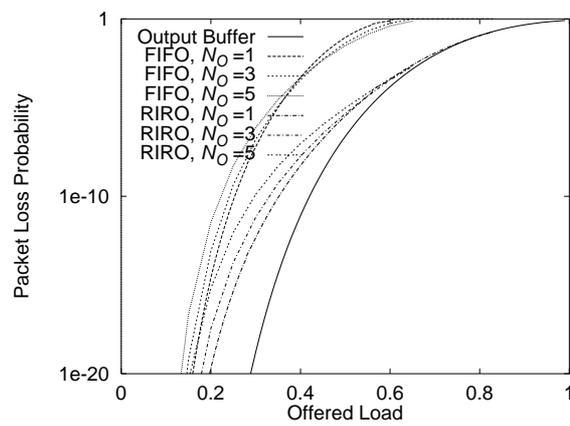


Figure 19: Packet Loss Probability vs. Offered Load for $\overline{BL} = 3$.

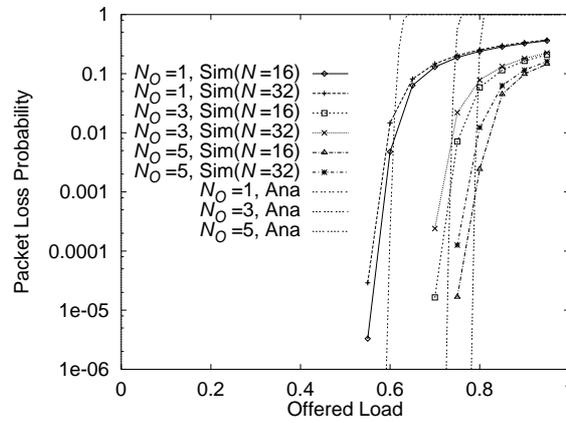


Figure 20: Comparison with Simulation Results for $\overline{BL} = 1$.

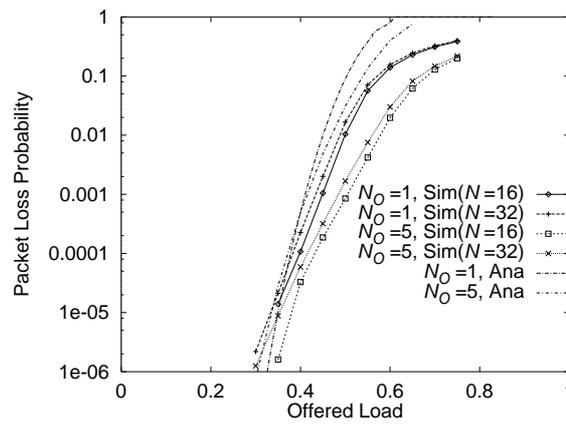


Figure 21: Comparison with Simulation Results for $\overline{BL} = 3$.

providing large output buffers. At least, the output buffer size comparable to the average packet length is necessary to gain a sufficient performance.

Recently, congestion control schemes such as the rate-based congestion control for ABR service class and EPD (Early Packet Discard) for UBR service class have been actively studied by many researchers [14, 15]. In most of their studies, the switch architecture is assumed to be ideal. That is, the internal switch speed is enough high so that congestion occurs at the output buffer. Thus, a threshold value associated with a single queue is considered to detect congestion. However, to implement these congestion control schemes in actual, performance limitations caused by the switch architecture should be taken into account as we have discussed in the current paper. For this purpose, our analytic results obtained in this paper can give the basis to investigate the congestion control mechanism in the ATM layer.

We further note that our analytic approach described in the current paper can be applied to the other cases, for example, the case where the switching speed is L ($1 \leq L \leq N$) times faster than the link speed (see, e.g., [16]), or the case where when $L' (> L)$ cells are simultaneously destined for the same output buffer, $(L' - L)$ cells are lost or kept awaiting at the input buffer.

For further works, we should evaluate the performance of the network in which two or more ATM switches are interconnected. In such a network, even when a long term congestion introduces large queue length at the input buffer, cell losses may be avoided to send back-pressure signals to the upper adjacent switches.

Acknowledgment

We would like to thank Dr. Hiroshi Suzuki and Dr. Ruixue Fan with NEC Corporation, C&C System Laboratories, for their invaluable suggestions.

References

- [1] P. Newman, "Traffic management for ATM local area networks," *IEEE Communications Magazine*, pp. 44–50, August 1994.

- [2] J. S. Turner, "Design of a broadcast packet switching network," *IEEE Transactions on Communications*, vol. 36, pp. 734–743, June 1988.
- [3] R. Rooholamini and V. Cherkassky, "Finding the right ATM switch for the market," *IEEE Computer*, pp. 17–28, April 1994.
- [4] R. Fan, H. Suzuki, K. Yamada, and N. Matsuura, "Expandable ATOM switch architecture (XATOM) for ATM lans," *ICC '94*, 5 1994.
- [5] I. Iliadis, "Performance of a packet switch with input and output queueing under unbalanced traffic," in *Proceedings of IEEE INFOCOM '92*, vol. 2, (Florence, Italy), pp. 743–752 (5D.4), 5 1992.
- [6] I. Iliadis, "Head of the line arbitration of packet switches with input and output queueing," in *Fourth International Conference on Data Communication Systems and their Performance*, (Barcelona, Spain), pp. 85–98, 6 1990.
- [7] I. Iliadis, "Synchronous versus asynchronous operation of a packet switch with combined input and output queueing," *Performance Evaluation*, no. 16, pp. 241–250, 1992.
- [8] A. I. Elwalid and I. Widjaja, "Efficient analysis of buffered multistage switching networks under bursty traffic," *IEEE GLOBECOM '93*, vol. 2, pp. 1072–1078, November–December 1993.
- [9] S. Gianatti and A. Pattavina, "Performance analysis of shared-buffered banyan networks under arbitrary traffic patterns," in *Proceedings of IEEE INFOCOM '93*, vol. 3, pp. 943–952, IEEE Computer Society Press, 3 1993.
- [10] M. J. Karol, M. G. Hluchyj, and S. P. Morgan, "Input vs. output queueing on a space-division packet switch," in *Proceedings of IEEE GLOBECOM '86*, (Houston, Texas), pp. 659–665, 12 1986.
- [11] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C*. Cambridge University Press, 1988.
- [12] D. Bertsekas and R. Gallager, *Data Networks*. Englewood Cliffs, New Jersey: Prentice-Hall, 1987.
- [13] H. Takagi, "Queueing analysis volume 3: Discrete-time systems," *North-Holland*, 1993.

- [14] H. Ohsaki, M. Murata, H. Suzuki, C. Ikeda, and H. Miyahara, "Rate-based congestion control for ATM networks," *ACM SIGCOMM Computer Communication Review*, vol. 25, pp. 60–72, April 1995.
- [15] H. Ohsaki, G. Hasegawa, M. Murata, and H. Miyahara, "Parameter tuning of rate-based congestion control algorithms and its application to TCP over ABR," *First Workshop on ATM Traffic Management IFIP WG 6.2*, pp. 383–390, December 1995.
- [16] Y. Oie, M. Murata, K. Kubota, and H. Miyahara, "Performance analysis of nonblocking packet switches with input / output buffers," *IEEE Transactions on Communications*, vol. 40, pp. 1294–1297, 8 1992.